

ZAI-SP 建模指南

文档版本：1.1

最后修改时间：2019-2

作者：qq600585

目录

什么是 SP?	2
在 zAI 中使用 SP 工作需要什么条件?	2
SP 技术指标.....	2
SP 在人脸识别中的作用	3
SP 在换脸应用中的作用	3
SP 在工业生产中的作用	3
SP 建模的准备工作 and 制作思路.....	4
解释：为什么 SP 搭配 OD 是最佳组合，而不是 SP 搭配 MMOD	5
开始制作我们的 OD.....	6
开始训练我们的 OD.....	7
以我们刚才训练的 sp_od_1.svm_od 从新生成框体	8
干掉没有被自动标记出来的框体.....	9
开始在 OD 框体上定义 SP 几何形.....	10
依次按顺序，去描绘所有框体内的图像.....	11
开始训练 SP.....	12
使用刚才我们训练的小模型开始制作大数据集.....	13
合并数据集.....	15
使用工具将合并数据集标记出 sp 坐标	16
小范围做一下 sp 编辑矫正	17
迭代式训练 SP.....	18
大规模训练非常耗时.....	18
使用 Remote file 查看远程服务器的训练状态	19
最后，我们自己的 14 坐标面部模型制作完成.....	21

什么是 SP?

SP: Shape Predictor 亦或 shape Detector, 顾名思义, 它是一种检测图像内容, 并将内容以几何方式输出。而几何图形由坐标组成, 所以 SP 对图像的处理就是检测内容, 输出坐标, 至于坐标可以组成什么样的几何形, 那就是 SP 建模的工作了。

在 zAI 中使用 SP 工作需要什么条件?

产生几何形坐标点, 只需要图片, 就是像素。处于优化的考虑, zAI 中的 SP 是工作于剪裁图像的, 比如, 我们使用 svm_od 检测到了一个物体, 这时候, 就会得到一个框体(rect), 而框体(rect), 它就是一副剪裁过的图像, SP 会把像素分析的焦点对准剪裁图像。

最简单的解释, SP 需要一幅图像, 一个图像内的 Rect, 这时候, 它就会从图像中分析出几何图形: 点坐标。

SP 技术指标

运行平台: 任何平台

并行: 支持

运行时内存消耗: 中

运行时性能: 高

CPU 需求: 高

GPU 需求: 无

超参数复杂度: 傻瓜化

训练后的文件尺寸: 非常大

次要指标

残差网络支持: 不支持

光照变换训练: 不支持

深度学习支持: 不支持

训练内存消耗: 大规模训练内存消耗会巨大无比, 小规模训练它会很快完成

训练时长: 小规模很快, 大规模很久

能支持的最大数据集训练: 我最大测试了 13000 张图片, cpu 训练实在太慢了

SP 在人脸识别中的作用

SP 在人脸识别中，主要用于预处理环节，它会检测出面部五官的几何构成，然后做对齐，对齐其实就是一种投影，因为有了几何坐标系，可以知道侧面到正面的投影参数。通过投影，对齐人脸后，人脸就会都是正面，这样再 input net 去判断个人身份，提高准确率。ZAI 的人脸在 SP 环节是自动化处理的，我们无需再自行做投影对齐

SP 在换脸应用中的作用

变脸 app 很多人都玩过，它能在你的头上戴上帽子，把你打扮的圣诞老人，变脸 app 的检测技术就是 sp，其原理就是当它定位了五官坐标以后，在五官坐标上画出覆盖的图形

SP 在工业生产中的作用

因为 SP 将图像描述成了几何形状，因为，我们通过分析 SP 的平面坐标，可以知道改刀，扳手，它们的手柄在哪里。

通过 SP，我们能准确判断出螺丝转动了多少圈，多少度

通过 SP，我们能对图像进行再次投影和裁剪，我们可以先检测机器轮廓(OD+SP)，然后再检测机器的部件(用 sp 描述部件形状)

SP 建模的准备工作和制作思路

本节为高级内容，研究本节内容，请先基于初级的 **ZAI-OD 建模指南.pdf** 实际动手一次

重要提醒：SP 建模需要在 **svm_od** 的基础展开，如果我们要做手机的几何构成图识别，我们就需要先做手机的 **svm_od**（OD 技术体系，使用 **z_ai_model** 可以将手机数据集训练成 **svm_od**）。**svm_od** 的生成请参考文档 **ZAI-OD 建模指南.pdf**

本节是高级内容，这里只会指出制作步骤，细节就不去一一解释了

第一步，准备足够多的识别目标样本

第二步，将目标样本分为两批，A 是少量的高质量样本（10 张图），B 是大批量样本（假设 1000 张图以上）

第三步，使用 OD 工艺手工制作框体 A 样本（**ZAI-OD 建模指南.pdf**）

第四步，将 A 样本训练成 **svm_od**（**ZAI-OD 建模指南.pdf**）

第五步，使用 A 样本训练出来的 **svm_od**，自动化框住 A 样本对象（不要用原始的框体定义，要用 **svm_od** 来框住自己）

第六步，在 A 样本的 **svm_od** 自动框体上，编辑 SP 几何形

第七步，将 A 样本的 SP 几何形训练成 **shape**

第八步，将 A 样本的 **svm_od** 和 **shape**，自动应用于 B 样本

第九步，调教将 B 样本，并且将 B 样本训练成 **svm_od+shape**，如此反复迭代

解释：为什么 SP 搭配 OD 是最佳组合，而不是 SP 搭配 MMOD

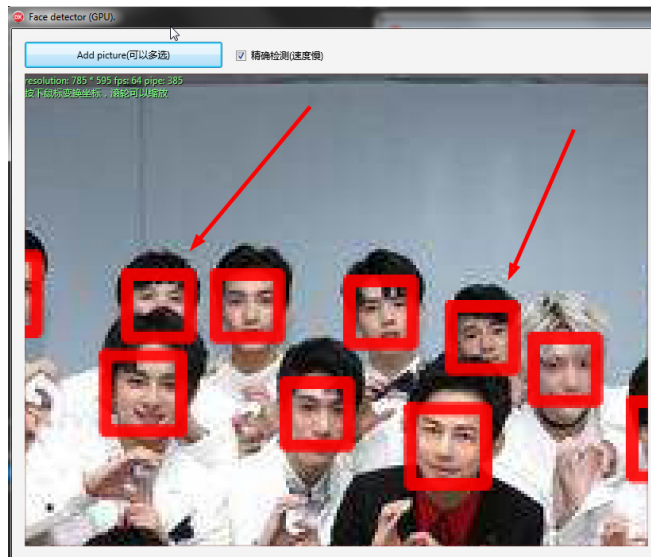
OD 检测出来的是完整目标图像，且使用和训练都非常简单，非常易于上手使用

MMOD 太过于智能化，它能检测出来重叠目标，甚至变形目标，而重叠变形目标的图像信息往往是丢失的，以 zAI 的 Demo 为例

下图的用 OD 技术检测人脸 Demo，箭头指向因为遮挡，导致信息不全，没有被检测到



下图是 MMOD 技术检查人脸 Demo，MMOD 使用 resnet 残差网络训练，所以即使遮挡它能检测出来

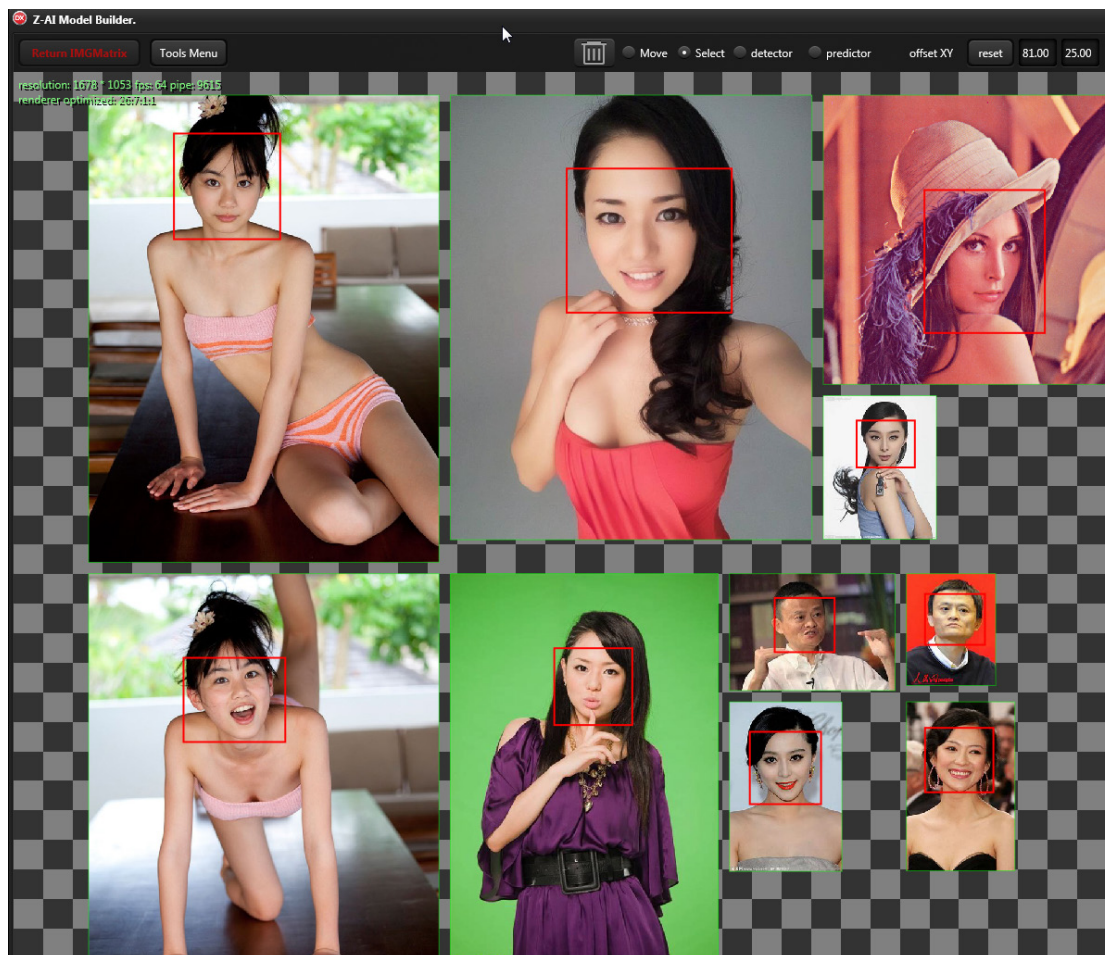


被遮挡的人脸，对SP来说没有可用性¹

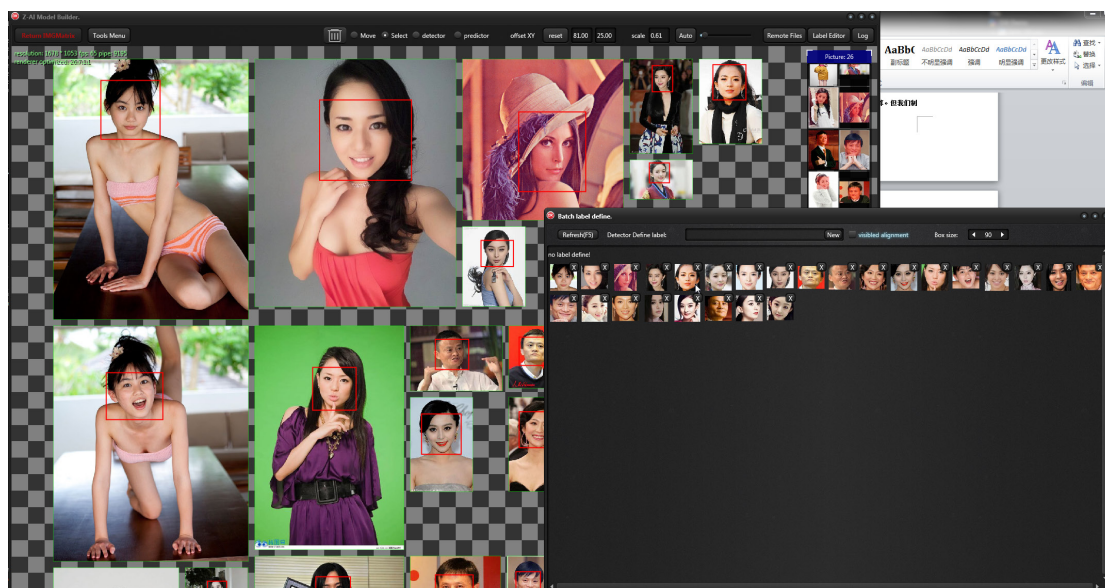
¹¹当信息不全的时，SP 计算就会没有可参考的数据，没有可参考数据，这时候 SP 会猜测去判断，虽然 Sp 也能计算。但我们制作 SP 模型不应该使用有重叠检测功能的 OD。误判将会导致 SP 的模型质量下降

开始制作我们的 OD

首先，我们框住需要检测的目标对象，这些框都是不准确的人工制作框体

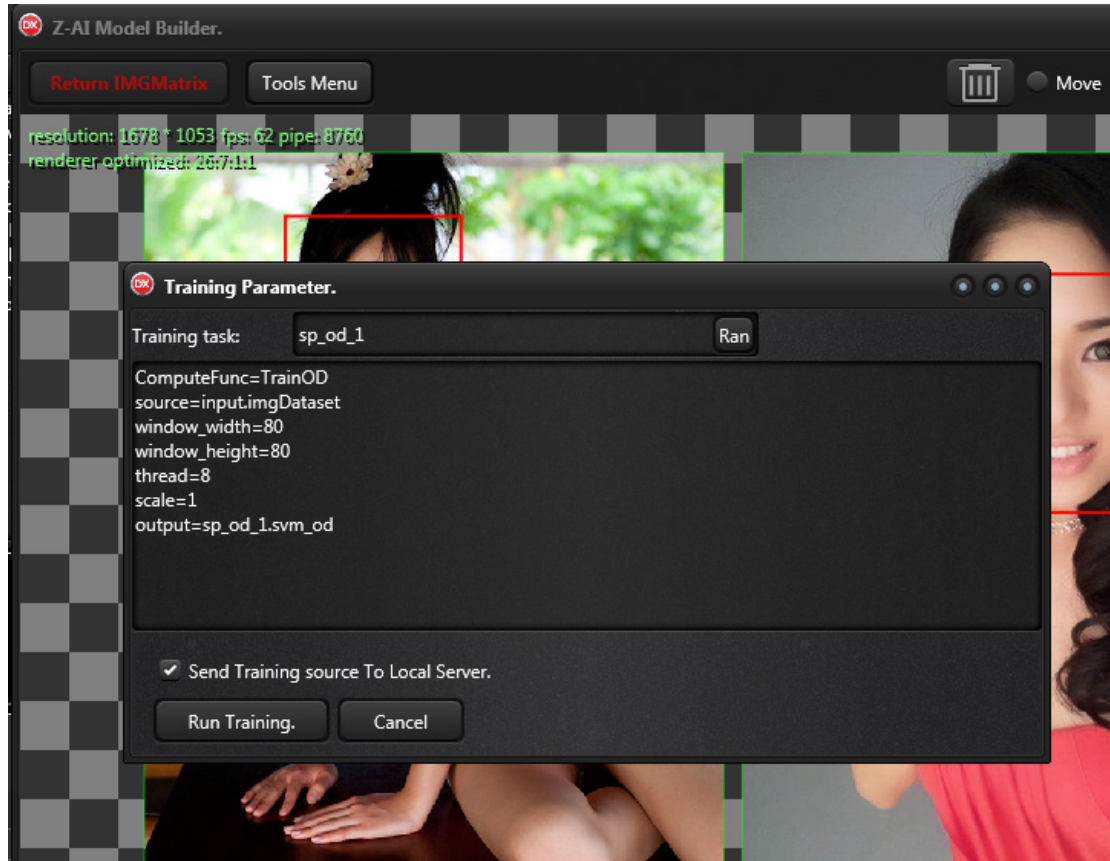


这时候全图如下，一共 26 个人头框体，都是手工拉框

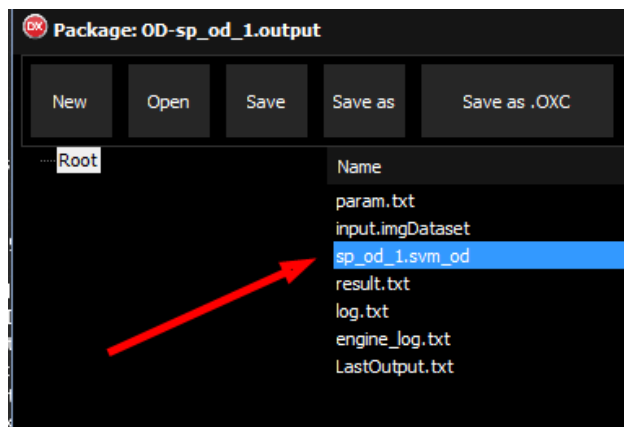


开始训练我们的 OD

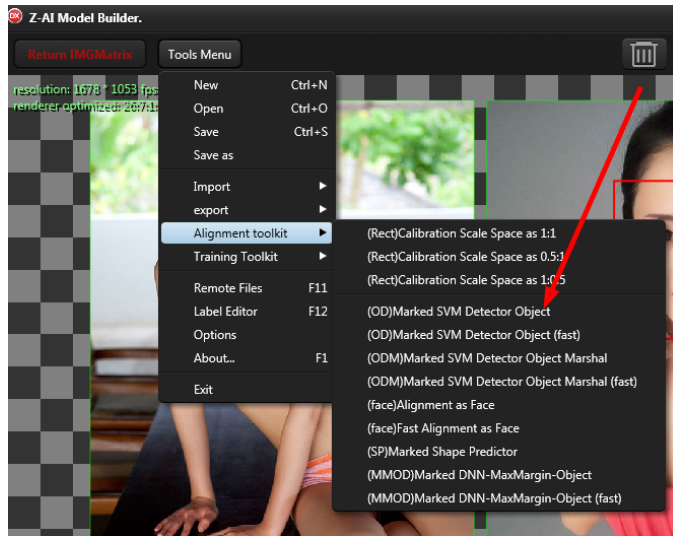
点 Run Training.



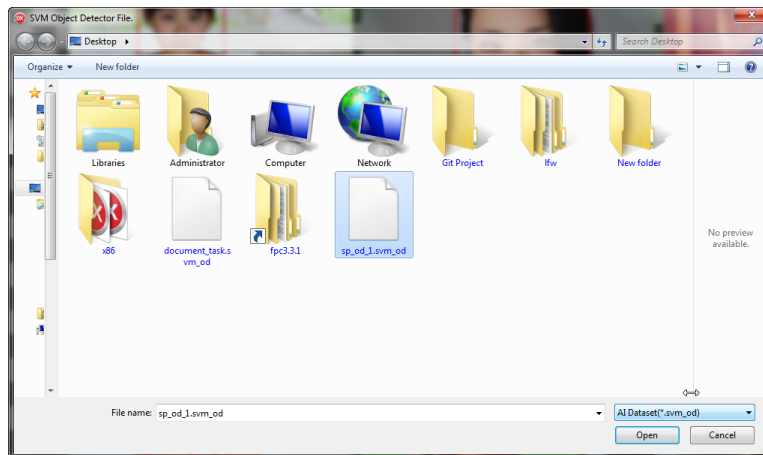
待训练完成，我们将 sp_od_1.svm_od 导出到桌面来



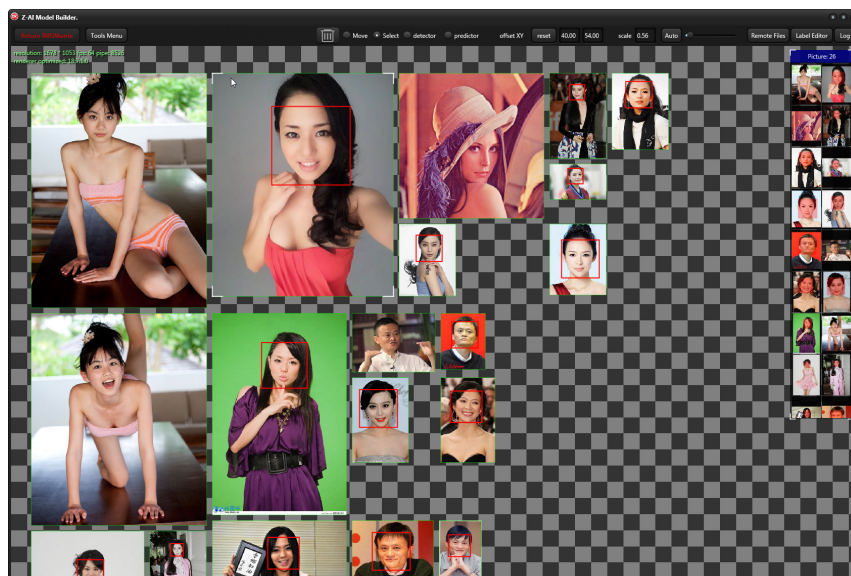
以我们刚才训练的 `sp_od_1.svm_od` 从新生成框体



选中 `sp_od_1.svm_od`

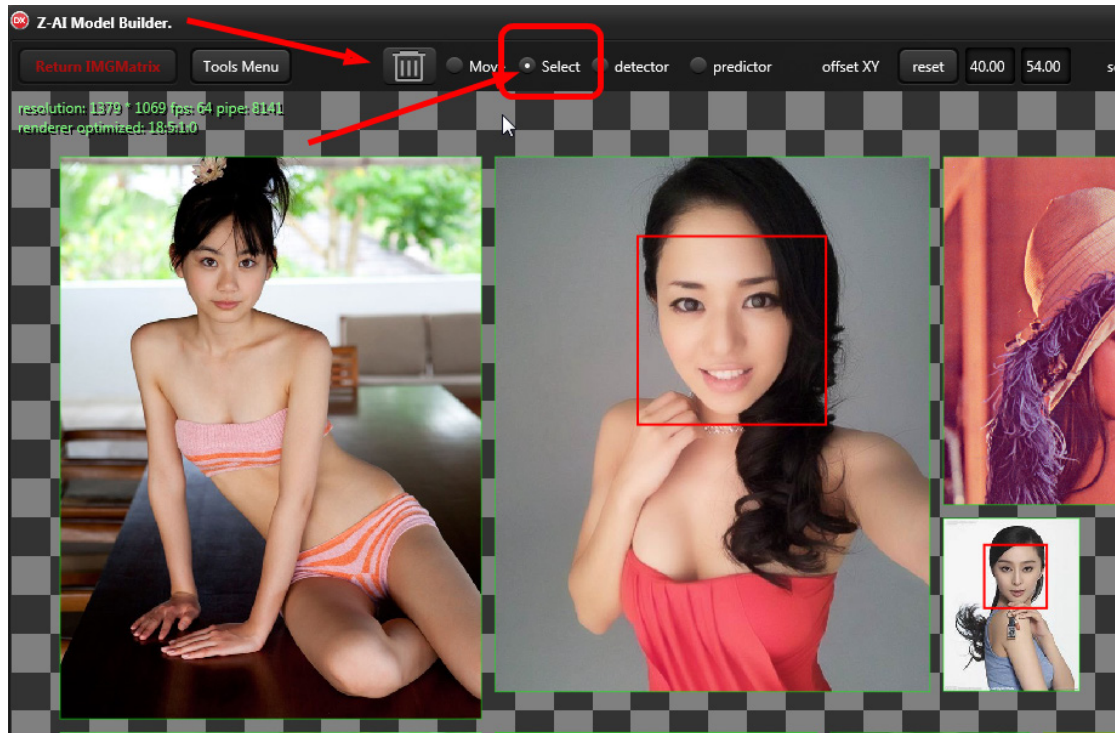


这时候，我们的目标就被重新定义框体了

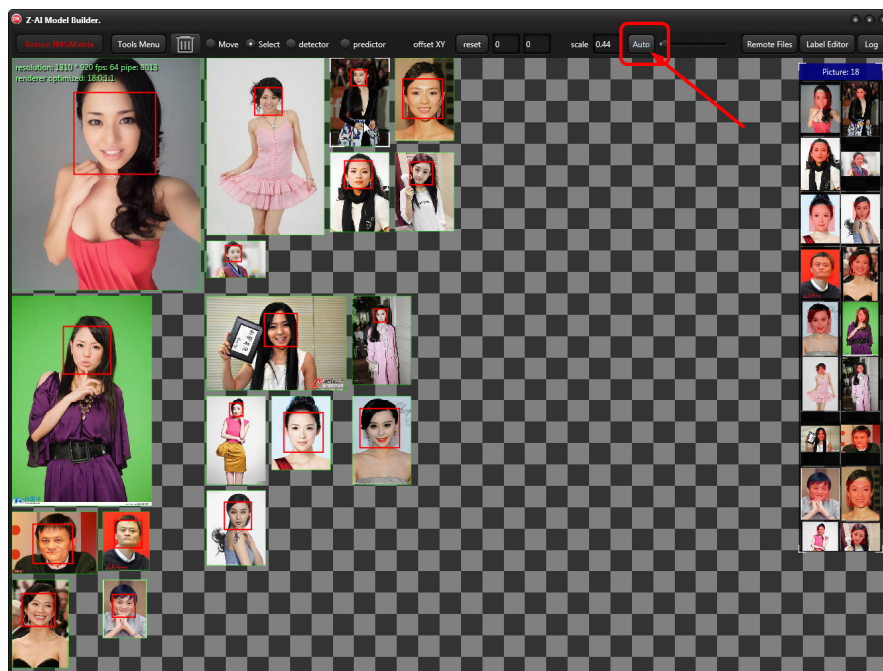


干掉没有被自动标记出来的框体

先选中图片，然后用 del 键或则点击垃圾桶来删除

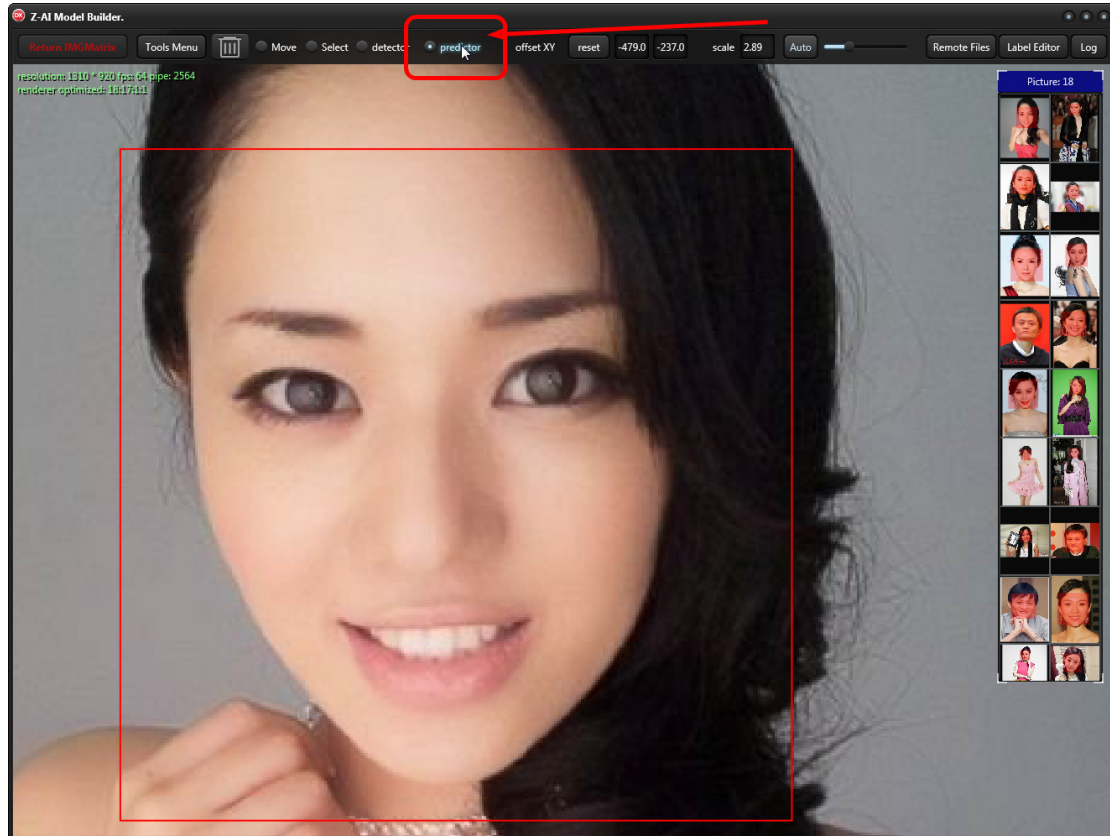


干掉以后，点一下 auto 重新排序，现在，都是刚才 od 检测出来的框体的图片了



开始在 OD 框体上定义 SP 几何形

切换至 predictor，然后将图片放大，一直到可以显示每个坐标的 index 提示



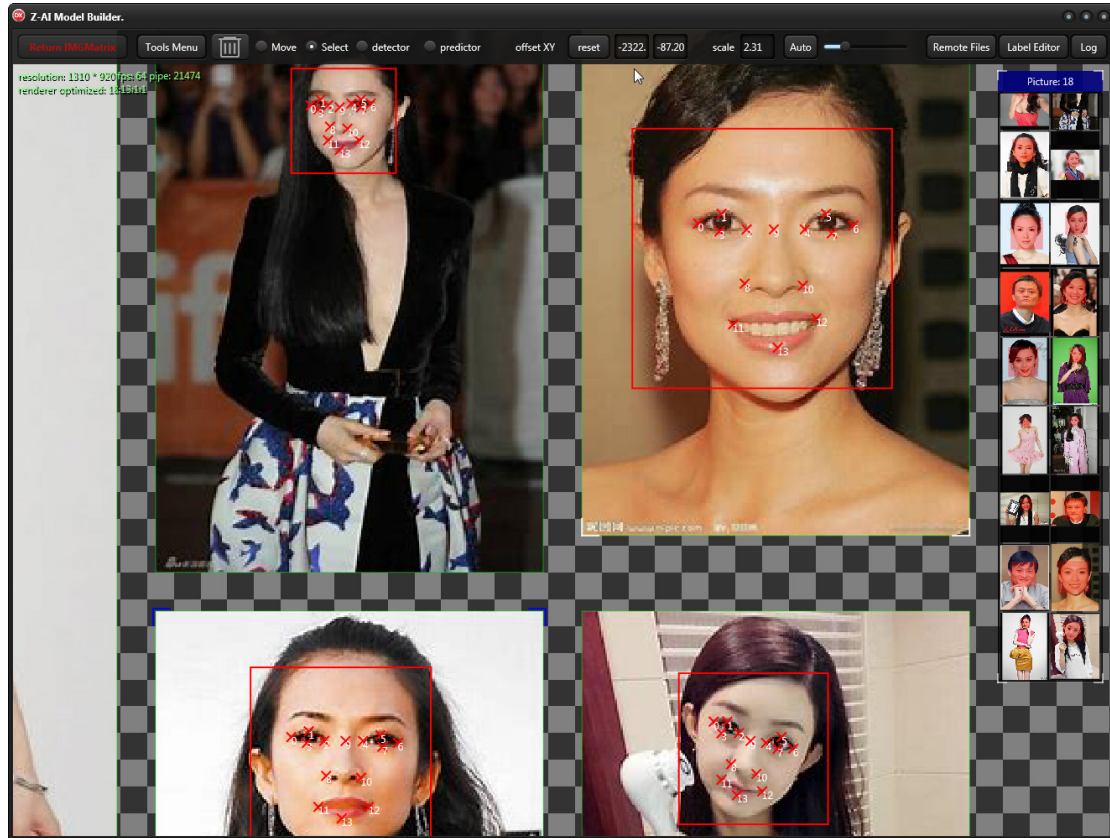
我们需要明确，我们需要将图像描绘成什么样的几何形，然后按顺序点，我将眼，鼻，嘴，定义成如下几何

0-3 顺时针描绘左眼，4-7 顺时针描绘右眼，8-10 顺时针描绘鼻子，11-13 顺时针描绘嘴

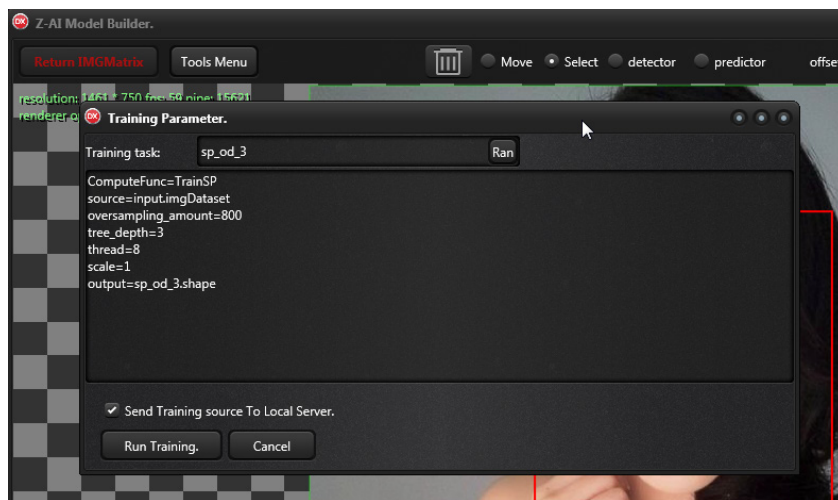
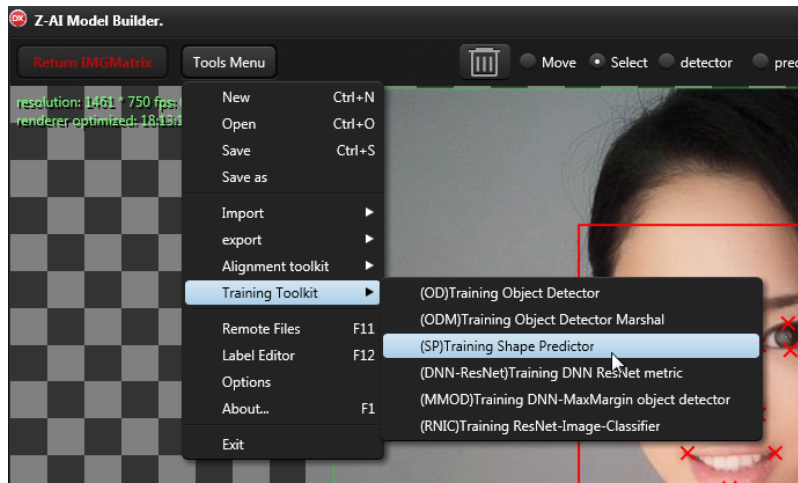


依次按顺序，去描绘所有框体内的图像

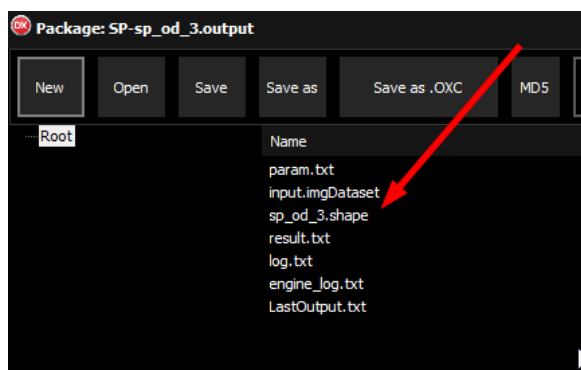
注意：全都按顺序来，本文定义的几何顺序结构如下



开始训练 SP



本文数据集很小，1 分钟以后训练完成，然后将 `sp_od_3.shape` 导出到桌面



使用刚才我们训练的小模型开始制作大数据集

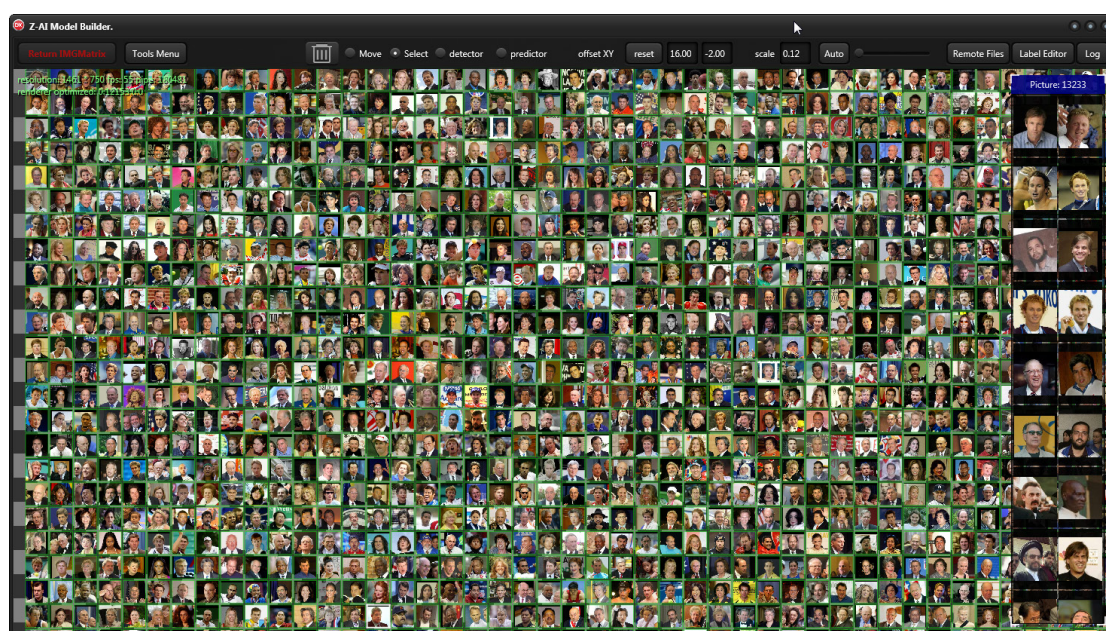
本节是迭代式训练的第二个大环节

阅读本节先要确定下面两个模型已经存在

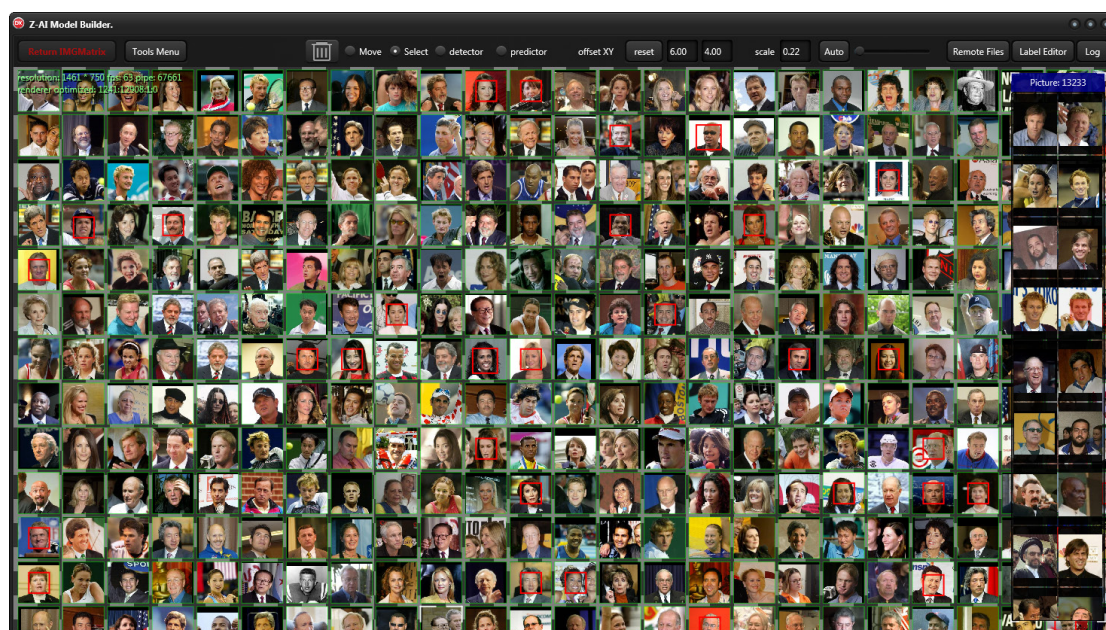
`sp_od_1.svm_od`

`sp_od_3.shape`

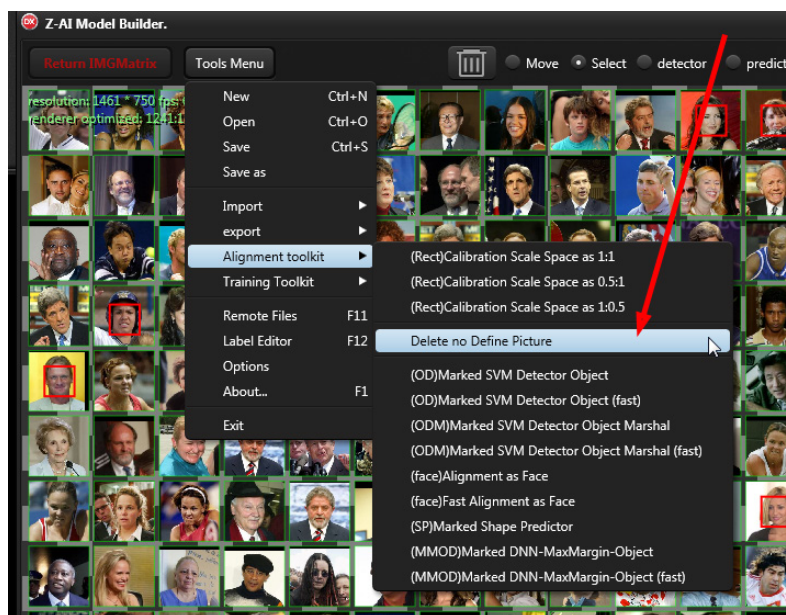
我们先使用 `z_ai_model` 工具导入一大批目标进来，总共 13000 张



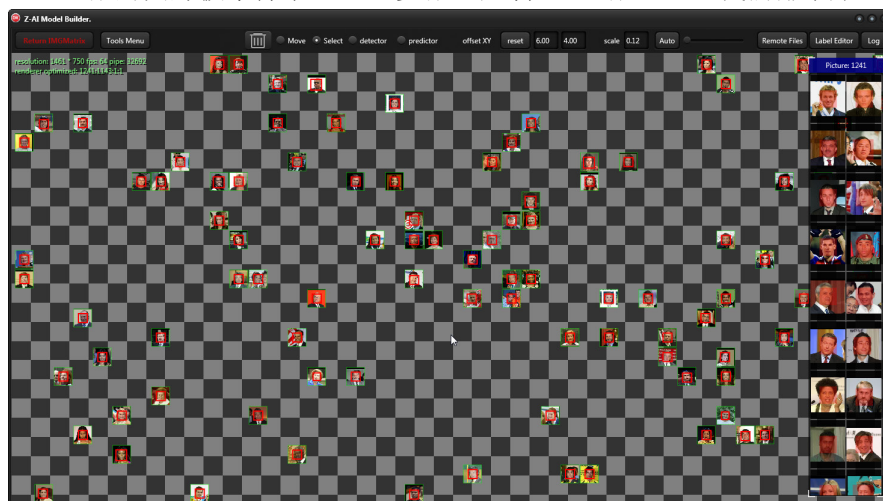
然后使用 `sp_od_1.svm_od` 将框体标记出来



这时候，很多图片没有被标记目标，我们干掉他们，使用 z_ai_model 批处理工具

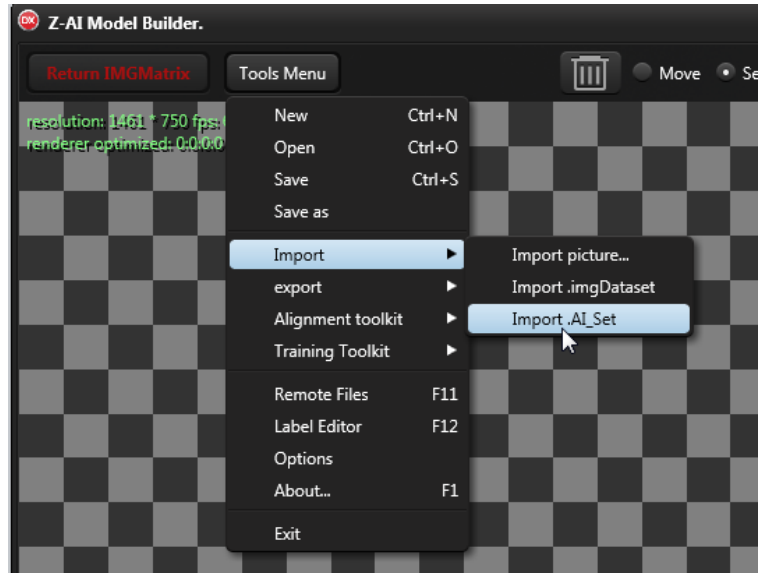


13000 张图片被干掉了 11000 多张，只剩 1200 张左右，我们排个序

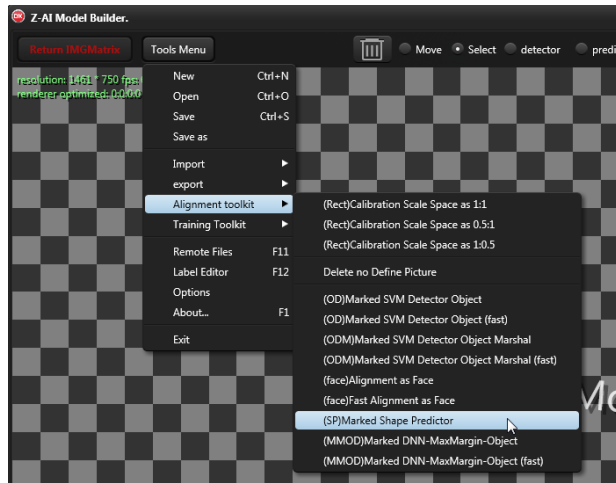


合并数据集

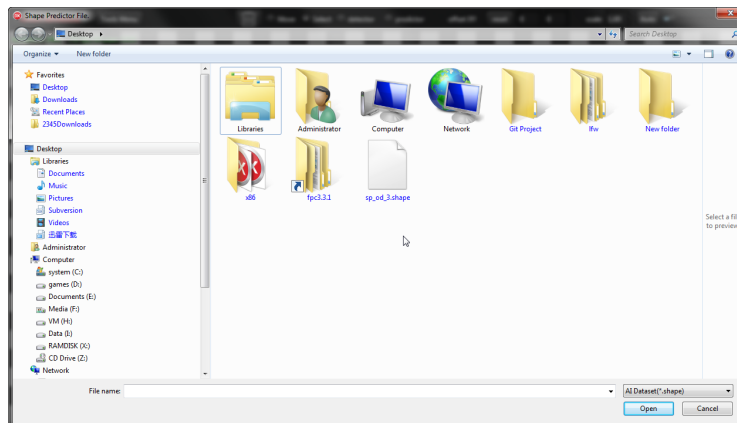
从 1241 到 1258，新增大概 17 张



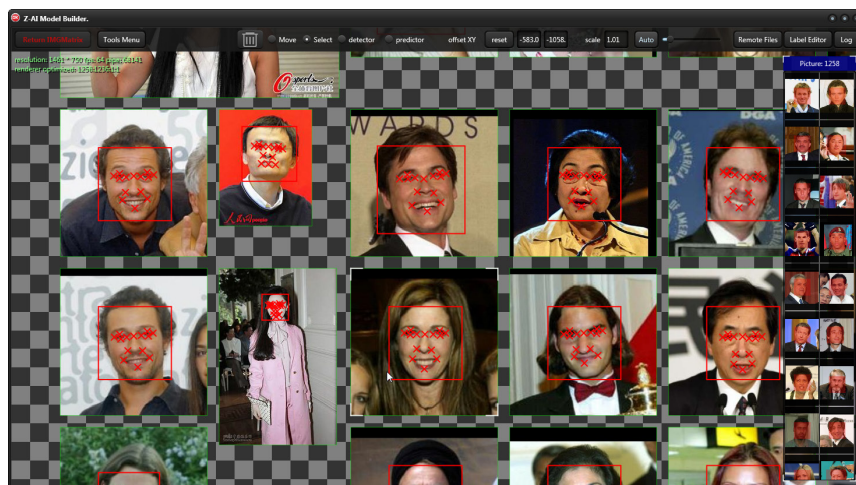
使用工具将合并数据集标记出 sp 坐标



选择我们刚才训练的小规模 sp

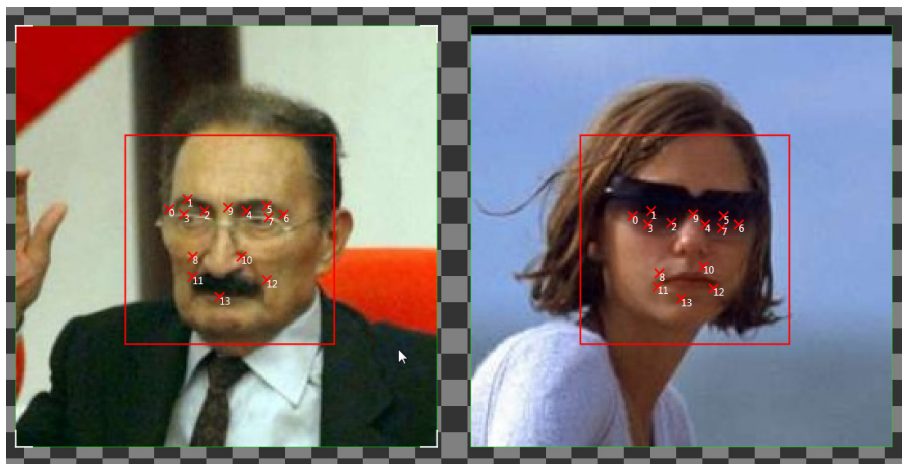


现在我们有 1258 张图片，完成自动框体+标记

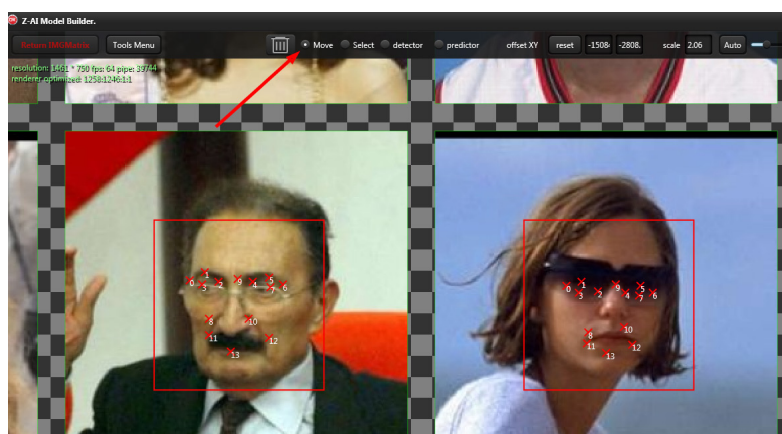


小范围做一下 sp 编辑矫正

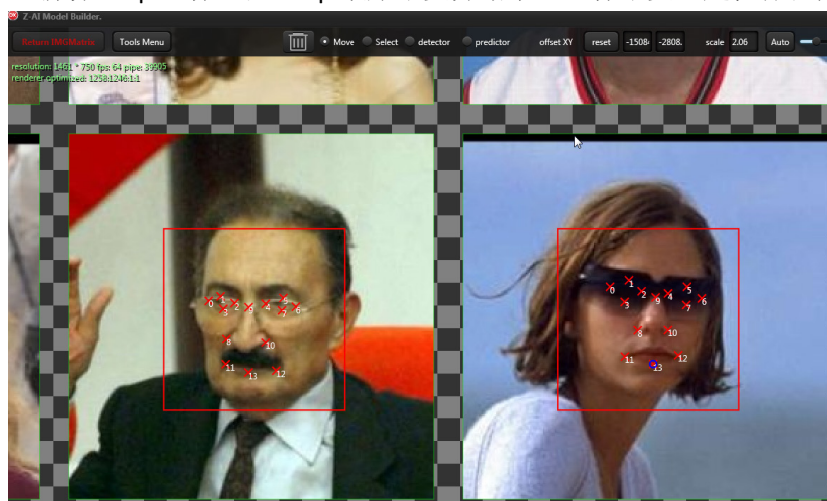
我们通过浏览，发现留蓄须和戴墨镜的 sp 有错误，这时候，我们需要调整一下 sp



切换至 move 操作



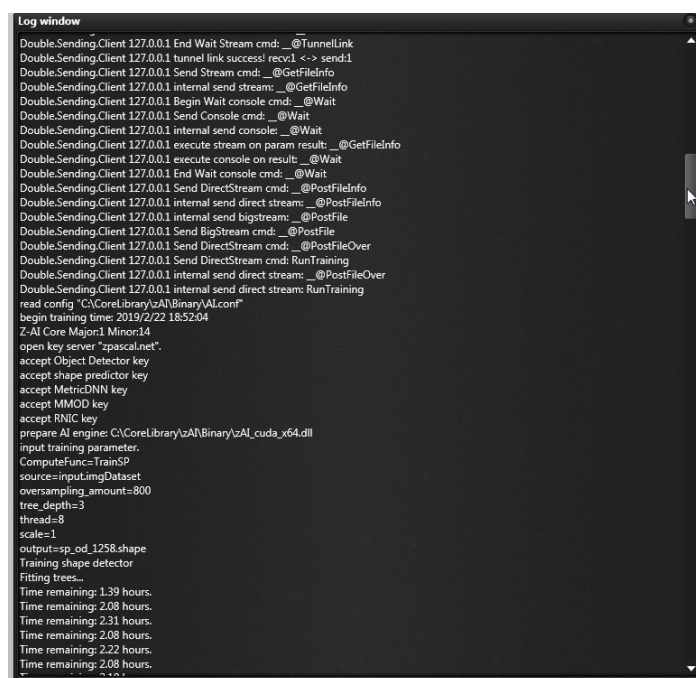
重新矫正 sp 坐标，注意 sp 顺序不要弄错了，坐标不要逾越框体范围



迭代式训练 SP

待所有的调整完成，存盘，再次训练本次做的 sp 修改

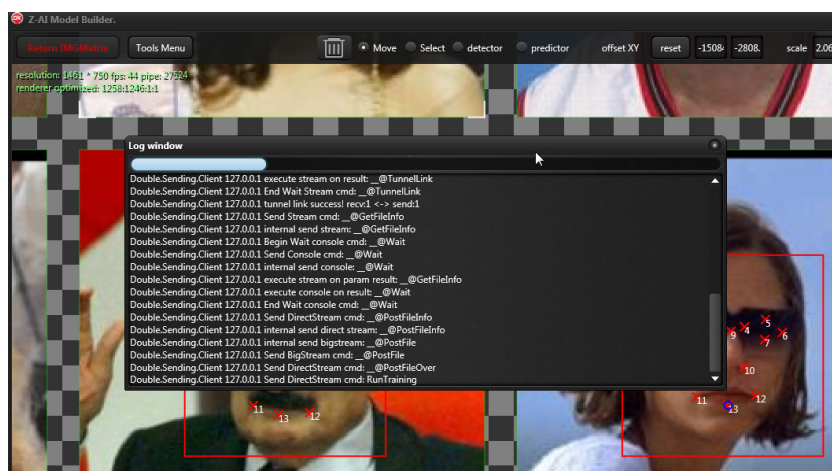
sp 训练是在 cpu 进行的。Sp 的图片越多，训练越久，耐心等待，本文在这里训练的图片 1300 张，大约 2 小时才能完成



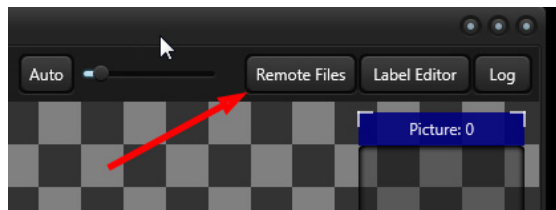
大规模训练非常耗时

我们待进度条走完就可以关闭 `model` 了，让服务器去慢慢训练。

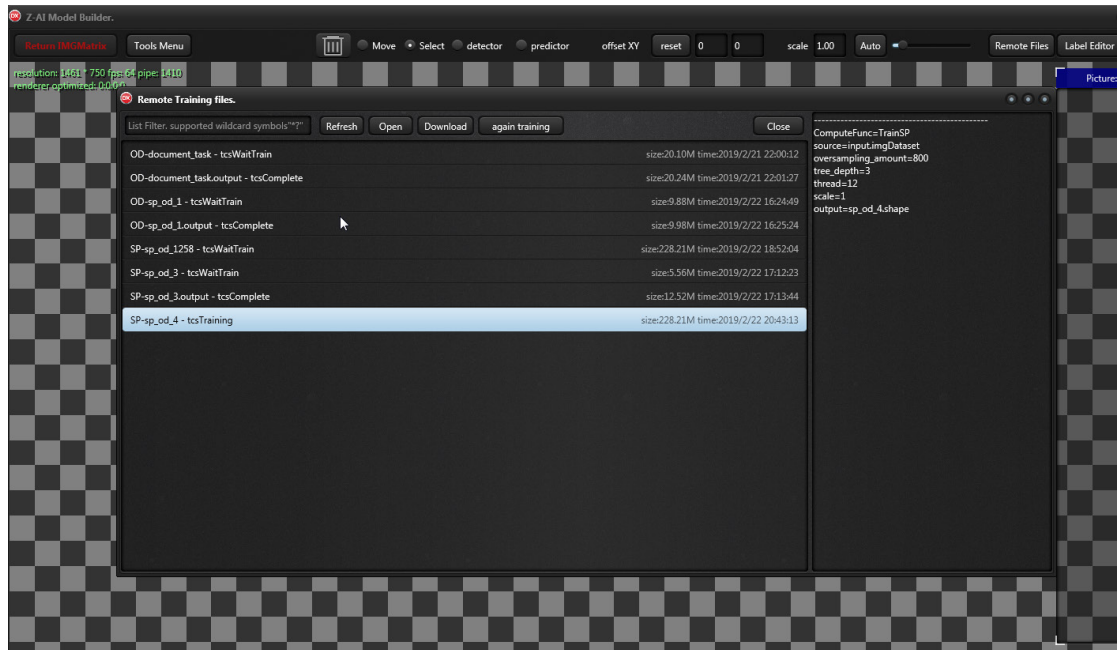
提示: `sp` 和 `od` 都是 `cpu` 在训练, 不会像 `gpu` 训练那样可以保存进度, 中间务必保证服务器不要宕机。



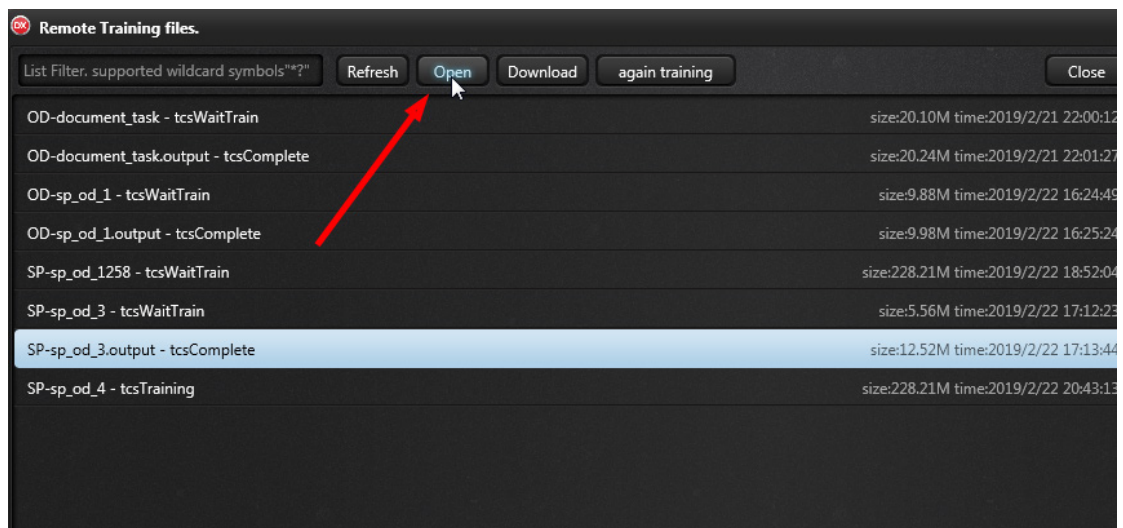
使用 Remote file 查看远程服务器的训练状态



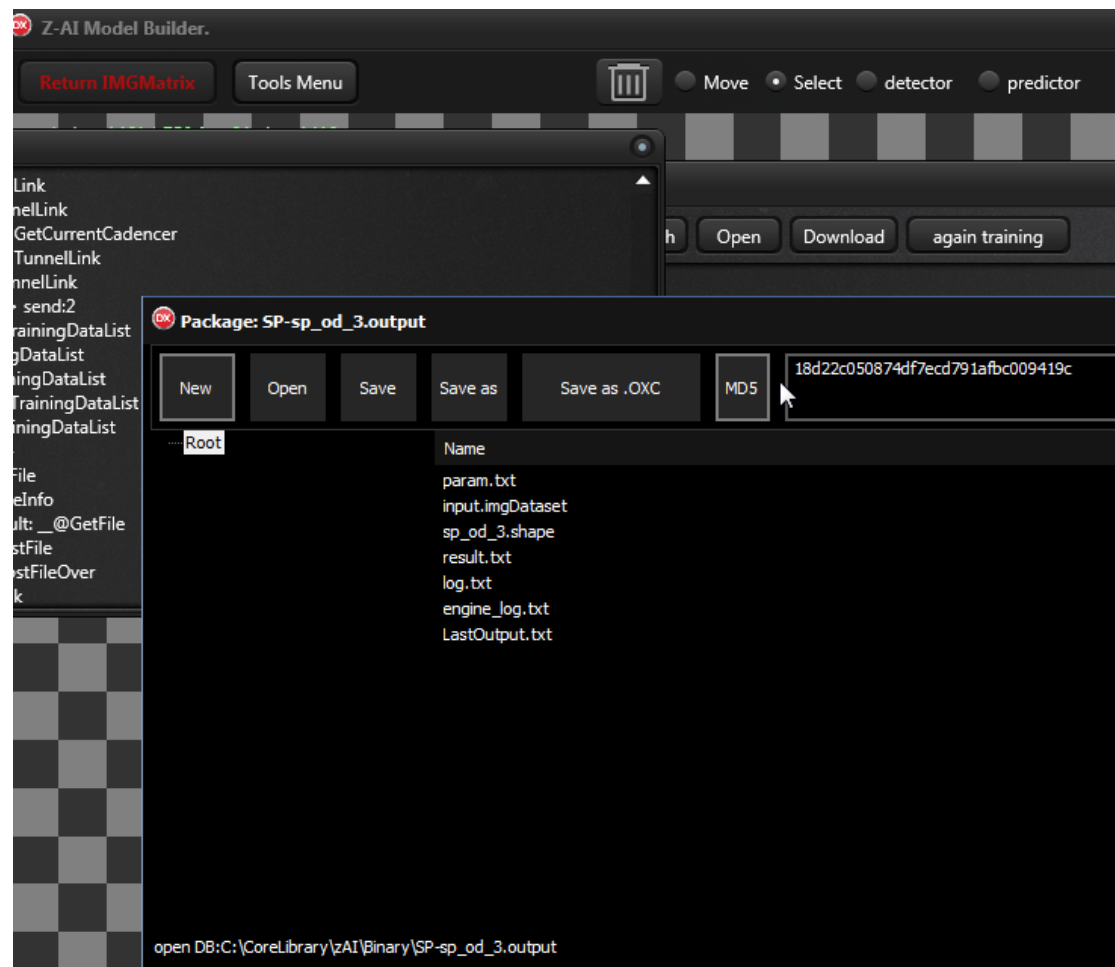
tcsTraining 表示正在训练中，等待训练完成



当训练完成后.点 Open，即可打开训练结果



将训练结果导出来即可



最后，我们自己的 **14** 坐标面部模型制作完成

照此工艺，不仅可以是制作人脸几何识别，同样也能制作物体几何识别。

对于 SP 的处理和应用请参考 zAI 的 Demo.内有详细说明。

如有使用问题，可以给我留言。

By.qq600585